Generating Generalized Inverse Gaussian Random Variates

Wolfgang Hörmann, Josef Leydold

Research Report Series
Report 123, February 2013

Institute for Statistics and Mathematics
http://statmath.wu.ac.at/

Generating Generalized Inverse Gaussian Random Variates

Wolfgang Hörmann · **Josef Leydold**

Abstract The generalized inverse Gaussian distribution has become quite popular in financial engineering. The most popular random variate generator is due to Dagpunar (1989). It is an acceptance-rejection algorithm method based on the Ratio-of-uniforms method. However, it is not uniformly fast as it has a prohibitive large rejection constant when the distribution is close to the gamma distribution. Recently some papers have discussed universal methods that are suitable for this distribution. However, these methods require an expensive setup and are therefore not suitable for the varying parameter case which occurs in, e.g., Gibbs sampling. In this paper we analyze the performance of Dagpunar's algorithm and combine it with a new rejection method which ensures a uniformly fast generator. As its setup is rather short it is in particular suitable for the varying parameter case.

Keywords random variate generation · generalized inverse Gaussian distribution · varying parameters

Mathematics Subject Classification (2000) 65C05 · 65C10

1 Introduction

The *generalized inverse Gaussian* (GIG) distribution has become quite popular for modeling stock prices in financial mathematics (Eberlein and Keller 1995). This distribution was first proposed by Etienne Halphen and thus it is sometimes referred as Halphen's law, see

Paper accepted for publication in *Statistics and Computing*.

W. Hörmann
Department of Industrial Engineering
Boğaziçi University
34342 Bebek-İstanbul, Turkey
E-mail: hormannw@boun.edu.tr

J. Leydold
Institute for Statistics and Mathematics
WU (Vienna University of Economics and Business)
Augasse 2–6, A-1090 Wien, Austria
E-mail: josef.leydold@wu.ac.at

Seshadri (1999). Its name has been established by Barndorff-Nielsen et al (1978), and its statistical properties have been investigated by Jørgensen (1982), see also Johnson et al (1994, p. 284) for further remarks on the history of this distribution.

The density of the GIG distribution is given by

$$f_{\mathrm{gig}}(x|\lambda,\psi,\chi) = \begin{cases} \dfrac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\psi\chi})}\, x^{\lambda-1}\exp\left(-\dfrac{1}{2}\left(\dfrac{\chi}{x}+\psi x\right)\right), & x>0\,,\\ 0\,, & x\le 0\,, \end{cases}$$

where $K_\lambda(\cdot)$ denotes the modified Bessel function of the third kind with index λ (Jørgensen 1982; Abramowitz and Stegun 1972). Parameters λ, ψ, and χ have to satisfy the condition

$$\lambda\in\mathbb{R}, \qquad (\psi,\chi)\in\begin{cases}\{(\psi,\chi)\colon\ \psi>0,\chi\ge 0\}, & \text{if } \lambda>0\,,\\ \{(\psi,\chi)\colon\ \psi>0,\chi>0\}, & \text{if } \lambda=0\,,\\ \{(\psi,\chi)\colon\ \psi\ge 0,\chi>0\}, & \text{if } \lambda<0\,.\end{cases}$$

The gamma distribution is a special case of this family with $\chi=0$. Thus if we exclude this case we get an alternative parametrization by setting $\alpha=\sqrt{\psi/\chi}$ and $\beta=\sqrt{\psi\chi}$. Then α becomes a mere scaling parameter and thus it is sufficient to consider the two parameter family of distributions with quasi-density

$$g(x|\lambda,\beta) = x^{\lambda-1}\exp\left(-\frac{\beta}{2}\left(x+\frac{1}{x}\right)\right) \qquad \text{for } x>0, \text{ where } \beta>0. \tag{1}$$

We have neglected the normalization constant $1/(2K_\lambda(\beta))$ in the formula above as the random variate generation methods discussed in this paper do not require it. We also note that we only need to derive a generator for $\lambda\ge 0$, since $1/X$ is also a GIG distributed random variate when X is GIG distributed, but with λ replaced by $-\lambda$. Thus we assume throughout the paper that $\lambda\ge 0$. The mode of g is given by

$$m = m(\lambda,\beta) = \frac{\sqrt{(1-\lambda)^2+\beta^2}-(1-\lambda)}{\beta} = \frac{\beta}{\sqrt{(1-\lambda)^2+\beta^2}+(1-\lambda)} \tag{2}$$

where the second formula is computed by means of the inverse value of the mode of $g(1/x)$. Notice that for $\lambda<1$ and very small values of β the second formula is robust against round-off errors, in opposition to the first formula which may result in severe cancellation errors. In particular, when using floating point numbers then $(1-\lambda)^2+\beta^2$ and $(1-\lambda)^2$ do not differ for sufficiently small β ($<10^{-8}$) and we obtain $m=0$ which cannot be used for further computations.

Despite the increasing popularity of the GIG distribution only a few methods have been proposed in the literature to generate variates from this distribution. The most popular generator is based on the Ratio-of-Uniforms method with minimal bounding rectangle and has been independently proposed by Dagpunar (1989) and Lehner (1989). It is a quite simple algorithm. It works well when $\lambda\ge 1$ or $\beta\ge 0.5$. However, it does not have a uniformly bounded rejection constant. Indeed when $\lambda<1$ its performance deteriorates rapidly when β is close to 0. We give an analysis of its performance in Section 3 below.

Atkinson (1982) proposes a rejection algorithm with a double-exponential hat. However, it requires to solve a bivariate optimization problem numerically. Dagpunar (2007, Sect. 4.8) proposed an algorithm that is based on rejection from a gamma-distributed hat. It has a quite fast setup but its rejection constant is only acceptable if $\beta<\lambda$.

Recently, Leydold and Hörmann (2011) have shown that a method that is based on Gauss-Lobatto integration and Newton interpolation is suitable for sampling from the GIG distribution by numerical approximation, see Derflinger et al (2010) for details. The approximation error can be selected close to machine precision in the IEEE 754 double floating point format (i.e., $2^{-52} \approx 2 \times 10^{-16}$). As the GIG distribution is log-concave for $\lambda \geq 1$, see below, other universal methods, such as Transformed Density Rejection (Hörmann et al 2004, Chap. 4), or rejection from a piecewise constant hat ("Ahrens's method", see Ahrens (1995) and Hörmann et al (2004, Sect. 5.1)) are suitable when an exact rejection algorithm is required. These algorithms have fast generation times. We also observed in our experiments that in particular this numerical inversion method as well as Ahren's method are numerically robust within a huge parameter range. However, the setup for these methods is expensive. So their usage is not recommended in the varying parameter case.

In this paper we propose a new generation method for the difficult case $\lambda < 1$, $\beta < 0.5$. It has a uniformly bounded rejection constant and a fast setup. Thus the method is especially useful in the varying parameter case. The paper is organized as follows: We collect general results and definitions from the literature useful to characterize the performance of the Ratio-of-Uniforms method in Section 2. In Section 3 we present our results that characterize the performance of the Ratio-of-Uniforms methods for the GIG distribution as proposed by Dagpunar (1989) and Lehner (1989). We discuss the variants and numerical details of the calculation of the bounding rectangle in Section 4, whereas Section 5 presents our results on the variant of the Ratio-of-Uniforms method without mode shift. Section 6 presents a hat with uniformly bounded rejection constant while the corresponding algorithm is presented in Section 7.

2 The Ratio-of-Uniforms Method for Arbitrary Distributions

In this section we briefly review the Ratio-of-Uniforms method with shift parameter μ and minimal bounding rectangle introduced by Kinderman and Monahan (1977); $f(x)$ denotes an arbitrary quasi-density.

1. Draw a random point (U,V) uniformly from the set

$$\mathscr{A}_f = \{(u,v)\colon 0 < v \leq \sqrt{f(u/v+\mu)}, u/v > \mu\}\ . \tag{3}$$

2. Return $X = U/V + \mu$.

The shift parameter μ is generally set to the mode m of f. In order to generate the point (U,V) rejection from the minimal bounded rectangle (MBR) for $\mathscr{A}_f$ is used. It is given by

$$\mathscr{R} = \{(u,v)\colon u^- \leq u \leq u^+,\ 0 \leq v \leq v^+\}$$

where

$$\begin{aligned} v^+ &= \sup_{x>0} \sqrt{f(x)} = \sqrt{f(m)} \\ u^- &= \inf_{x>0} (x-\mu)\sqrt{f(x)} \\ u^+ &= \sup_{x>0} (x-\mu)\sqrt{f(x)}\ . \end{aligned}$$

By a simple geometric argument we can partly characterize the performance of the Ratio-of-Uniforms method with mode shift μ equal to the mode of the distribution.

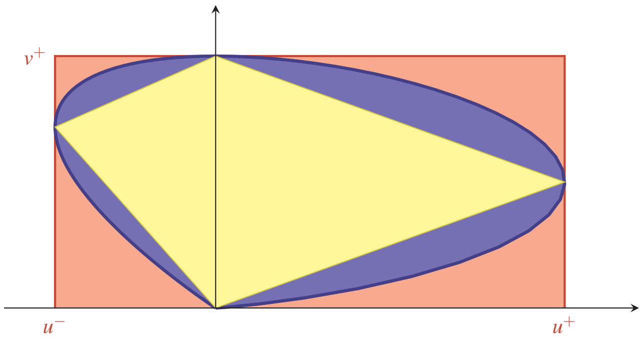

Fig. 1 Acceptance region $\mathscr{A}_f$, minimal bounding rectangle $\mathscr{R}$, and region of immediate acceptance.

Lemma 1 *If $\mathscr{A}_f$ is convex, then the rejection constant for rejection from the minimal bounded rectangle is bounded from above by* 2.

Proof Region $\mathscr{A}_f$ touches the minimal bounding rectangle $\mathscr{R}$ in the points $(0,0)$, $(u^+, u^+/x^+)$, $(0, v^+)$, and $(u^-, u^-/x^-)$. Thus the quadrangle $\mathscr{Q}$ with these vertices is contained in the convex set $\mathscr{A}_f$. Moreover, the area of $\mathscr{Q}$ is half of the area of $\mathscr{R}$, see Figure 1. Thus the result follows. □

For the sake of completeness we note that in the case where $\mathscr{A}_f$ is convex, quadrangle $\mathscr{Q}$ can be used as squeeze for the rejection algorithm. Moreover, it is then possible to apply the idea of immediate acceptance, see (Leydold 2000) for details. Then the marginal generation times decrease considerably at the expense of a slower setup.

To be able to check the condition of Lemma 1 directly from the quasi density the concept of T_c-concavity is very useful, see Hörmann (1995) for further details.

Definition 2 A function f is called T_c-concave ($c \neq 0$) if $\operatorname{sign}(c) f(x)^c$ is concave. It is called T_0-concave if it is log-concave.

Lemma 3 (Hörmann 1995) *If f is T_c-concave for some c, then it is $T_{c'}$-concave for every $c' \leq c$.*

The following result is crucial to characterize a large family of distributions that can be generated uniformly fast by the Ratio-uniforms-Methods.

Lemma 4 (Leydold 2000) *The set $\mathscr{A}_f$ is convex if and only if f is $T_{-1/2}$-concave.*

As a direct consequence of Lemmas 1 and 4 we obtain:

Corollary 5 *For a distribution with $T_{-1/2}$-concave density, the rejection constant for the ratio-of-uniforms method with mode shift and bounding rectangle is bounded from above by 2.*

3 Performance of the Ratio-of-Uniforms Method with Mode Shift

We return to the GIG distribution with quasi-density $g(x|\lambda, \beta)$ (defined in (1)). To find an upper bound for the rejection constant of the Ratio-of-Uniforms method it is important to identify the parameter region, where g is $T_{-1/2}$-concave. We therefore first investigate the concavity properties of the GIG quasi-density g. By a straightforward computation we obtain the following well-known result.

Lemma 6 *g is log-concave for $\lambda \geq 1$. Moreover, if $\lambda < 1$ then g is log-concave in $[0, x_0]$ and log-convex in $[x_0, \infty)$, where*

$$x_0 = \frac{\beta}{1-\lambda} \ . \tag{4}$$

Lemma 7 *Quasi-density $g(x|\lambda, \beta)$ is T_c-concave for $c = -\frac{1}{4\beta}$.*

Proof The statement holds for $\lambda \geq 1$ by Lemmata 3 and 6. So we have to show that $-g(x)^c = -x^{c(\lambda-1)}\, e^{-\frac{c\beta}{2}\left(x+\frac{1}{x}\right)} = -x^{c(\lambda-1)}\, e^{-\frac{1}{8}\left(x+\frac{1}{x}\right)}$ is concave when $\lambda < 1$. Thus we show that $\frac{\partial^2}{\partial x^2}\left(-x^a\, e^{\frac{1}{8}\left(x+\frac{1}{x}\right)}\right) \leq 0$ for all $a \in \mathbb{R}$ and $x > 0$. A straightforward computation gives

$$\frac{\partial^2}{\partial x^2}\left(-x^a\, e^{\frac{1}{8}\left(x+\frac{1}{x}\right)}\right) = p(x,a) \cdot x^{a-4} e^{\frac{1}{8}\left(x+\frac{1}{x}\right)}/64$$

where

$$p(x,a) = -(x^4 + 16ax^3 + (64a^2 - 64a - 2)x^2 + (16 - 16a)x + 1) \ .$$

Notice that for fixed x, $p(x,a)$ is maximal for $a = -\frac{x^2-4x-1}{8x}$ and hence

$$\max_{a\in\mathbb{R}} p(x,a) = -8x(x-1)^2 \leq 0 \qquad \text{for all } x \geq 0.$$

This completes the proof. $\square$

If we inspect the proof of Lemma 7 again, then we also find that $p(x,a) = 0$ if $a = 1/2$ and hence if $\lambda = 1 - 2\beta$. Consequently, g is not T_c-concave in some open interval around $x = 1$ if $c < -\frac{1}{4\beta}$. So our result in Lemma 7 is sharp.

Corollary 8 *Quasi-density g is $T_{-1/2}$-concave when $\lambda \geq 1$ or $\beta \geq 1/2$.*

Lemma 9 *Quasi-density g is $T_{-1/2}$-concave if $0 < \lambda < 1$ and $\beta \geq \frac{2}{3}\sqrt{1-\lambda}$.*

Proof We have to show that $\left(1/\sqrt{g(x)}\right)'' \geq 0$ for the given parameter range. Because of Corollary 8 we assume that $\frac{2}{3}\sqrt{1-\lambda} \leq \beta \leq 1/2$ and consequently $7/16 \leq \lambda < 1$. A straightforward computation yields

$$\left(1/\sqrt{g(x|\lambda,\beta)}\right)'' = p(x,\lambda,\beta) \cdot x^{-\frac{1}{2}(\lambda+7)} e^{\frac{1}{4}\beta\left(x+\frac{1}{x}\right)}/16$$

where

$$p(x,\lambda,\beta) = \beta^2 x^4 + 4\beta(1-\lambda)x^3 + \left(4(\lambda^2-1) - 2\beta^2\right)x^2 + 4\beta(\lambda+1)x + \beta^2 \ .$$

Since $x \geq 0$ and thus

$$\frac{\partial}{\partial\beta} p(x,\lambda,\beta) = 2\beta(x-1)^2(x+1)^2 + 4(1-\lambda)x^3 + 4(\lambda+1)x \geq 0$$

we find that $p(x,\lambda,\beta)$ is monotonically increasing in β and hence its minimum is located on the curve $\beta = \frac{2}{3}\sqrt{1-\lambda}$ which can equivalently be expressed as $\lambda = 1 - \frac{9}{4}\beta^2$. Thus

$$\begin{aligned} p(x,\lambda,\beta) &\geq p(x,\beta) := p(x, 1-9\beta^2/4, \beta) \\ &= \beta^2 x^4 + 9\beta^3 x^3 + (\tfrac{81}{4}\beta^4 - 20\beta^2)x^2 + (8\beta - 9\beta^2)x + \beta^2 \,. \end{aligned}$$

In order to show that $p(x,\beta)$ is non-negative notice that $p(0,\beta) = \beta^2 > 0$ and $\lim_{x\to\infty} p(x,\beta) = \infty$. Moreover, for $x \geq 0$ the slope $(p(x,\beta) - p(0,\beta))/x$ of the secant through $(0,\beta)$ and (x,β) has a unique minimum in $x_c = \frac{1}{6}\sqrt{9\beta + 80/3} - 3\beta$ and thus

$$\frac{p(x,\beta) - p(0,\beta)}{x} \geq -\left(243\beta^5 + \sqrt{27\beta^2 + 80}(3^{7/2}\beta^4 + 80\sqrt{3}\beta^2) - 1836\beta^3 - 288\beta\right)/36$$

which is non-negative for all $\beta \in [0, 1/2]$. Consequently, $p(x,\beta) \geq p(0,\beta) = \beta^2 > 0$, which completes the proof. $\square$

When we combine all our previous results we find a parameter range with uniformly bounded rejection constant for the Ratio-of-Uniforms method.

Theorem 10 *If g is $T_{-1/2}$-concave, then the rejection constant of the Ratio-of-Uniforms method with mode shift is at most 2. This is in particular the case if $\lambda \geq 1$ or $\beta \geq \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$.*

We have already mentioned in the introduction that the rejection constant of Dagpunar's algorithm is not uniformly bounded. It is possible to show that for $\lambda \in [0,1)$ the reciprocal of the rejection constant, i.e., the acceptance probability $|\mathscr{A}_g(\lambda,\beta)|/|\mathscr{R}(\lambda,\beta)|$ vanishes for β tending to 0.

Lemma 11 *Let $\lambda \in [0,1)$. Then*

$$\lim_{\beta\to 0} \frac{|\mathscr{A}_g(\lambda,\beta)|}{|\mathscr{R}(\lambda,\beta)|} = 0\,.$$

The proof is lengthy and therefore deferred to the appendix.

Before we proceed we remark that alternative generation methods for $T_{-1/2}$-concave densities can be used as well. In particular Algorithm TDR with transformation $T_c(x) = -1/\sqrt{x}$ is a good choice, see Hörmann (1995). When g is even log-concave we may use transformation $T_0(x) = \log(x)$ which has already been proposed by Devroye (1986, Sect. 8.2.6, p. 301). However, then one needs to find appropriate construction points for the hat. Fortunately, it is not necessary to compute these points with high accuracy in opposition to the Ratio-of-Uniforms method where x^+ and x^- have to be computed up to machine precision since otherwise u^+ and u^- are selected too small and $\mathscr{R}$ does not enclose $\mathscr{A}_g$.

We also remark that there exist methods where only the mode and the normalization constant of the density have to be known. Devroye (1984) proposes such an algorithm for log-concave distributions. An algorithm that works for all $T_{-1/2}$-concave densities can be found in (Leydold 2001). These algorithms seem to be well-suited for the varying parameter case. Unfortunately, we found that they are rather slow as they require the computation of the Bessel function which is quite expensive and the rejection constant is rather high.

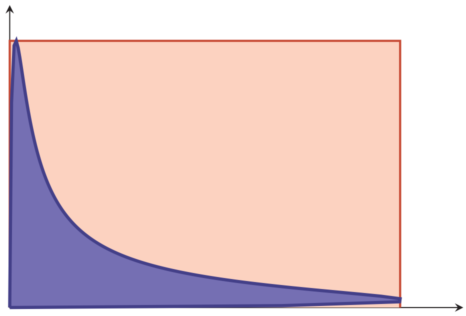

Fig. 2 Non-convex acceptance region $\mathscr{A}_g$.

4 Calculation of the Bounding Rectangle

Independently, Dagpunar (1989) and Lehner (1989) proposed generation algorithms for the GIG distribution based on the Ratio-of-Uniforms method with minimal bounding rectangle.

A straightforward computation shows (Dagpunar 1989) that u^- and u^+ are given by $(x^- - m)\sqrt{f(x^-)}$ and $(x^+ - m)\sqrt{f(x^+)}$, respectively, where x^- and x^+ are the two real roots of

$$x^3 - \left(m + \frac{2(\lambda+1)}{\beta}\right)x^2 + \left(\frac{2(\lambda-1)}{\beta}m - 1\right)x + m = 0 \tag{5}$$

which satisfy $0 < x^- < m$ and $m < x^+$, respectively. Dagpunar (1989) suggests a numerical method for finding these roots, in particular the bisection method, while Lehner (1989) makes use of Cardano's formula for solving this cubic equation (see Section "Solution of Cubic Equations" in Abramowitz and Stegun (1972)). We remark that neither Dagpunar (1989) nor Lehner (1989) provide any performance analysis for this algorithm.

Although the generator by Dagpunar (1989) is simple and very popular it has two serious drawbacks. The first one is the fact that the rejection constant diverges (see Lemma 11). In other words when the region $\mathscr{A}_g$ is non-convex, then the rejection constant may become very large, see Figure 2. Thus its performance decreases rapidly with $\beta \to 0$ when $\lambda < 1$; e.g., for $\lambda = 0.4$ and $\beta = 10^{-7}$ one finds a rejection constant of about 8500.

The second problem is the computation of the minimal bounding rectangle (MBR). It is quite expensive (for both approaches: numerical methods and Cardano's formula). Moreover, severe round-off errors occur when computing the MBR when $\lambda < 1$ and $\beta < 10^{-8}$. When Cardano's formula is used then one obtains `NaN` (Not-a-number). When a numerical root finding algorithm is used the situation is even worse: in all implementations that we tested in our experiments numeric problems result in much too small values for u^+ for that region of the parameter space. Therefore when $\beta \approx 10^{-8}$ the rejection constant drops from a very high value to 1 which obviously is much too small. This means that the generator silently returns random variables from a wrong distribution.

Due to these two problems we recommend not to use the RoU method with mode shift for the region with $\lambda < 1$ and $\beta < 0.5$.

5 Ratio-of-Uniforms without Mode Shift

Dagpunar (1988, Sect. 4.6.2) and Lehner (1989) also suggested to set $\mu = 0$ in (3), that is, not to shift the mode of quasi-density g into 0. Then equation (5) reduces to a quadratic equation and the boundaries of the minimal bounding rectangle are given by

$$\begin{aligned} v_0^+ &= \sup_{x>0} \sqrt{g(x)} = \sqrt{g(m)} \\ u_0^- &= \inf_{x>0} x\sqrt{g(x)} = 0 \\ u_0^+ &= \sup_{x>0} x\sqrt{g(x)} = x_0^+ \sqrt{g(x_0^+)} \end{aligned}$$

where x_0^+ is the positive root of (5) with $m = 0$ which is given by

$$x_0^+ = \frac{1}{\beta}\left((1+\lambda) + \sqrt{(1+\lambda)^2 + \beta^2}\right) . \tag{6}$$

Notice that we only have sums of positive numbers in this equations and no differences.

The important drawback of this method is that the rejection constant is not bounded from above. Indeed, computational experiences show that such an algorithm is useful only when both $\lambda \lessapprox 1$ and $\beta \lessapprox 1$ and one of the two parameters is close to 1. Nevertheless, it has the appealing property that computing (6) and thus the boundaries of the minimal bounding rectangle is much cheaper than solving the cubic equation (5). Thus it is quite attractive in the varying parameter case. The following results gives an upper bound for the rejection constant in a particular parameter range.

Lemma 12 *For* $0 \le \lambda \le 1$ *and* $\min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\} \le \beta \le 1$ *the rejection constant of the Ratio-of-Uniforms method without shift is bounded from above by* 2.619.

Proof Region $\mathscr{A}_g$ touches the bounding rectangle $\mathscr{R} = (0, u_0^+) \times (0, v_0^+)$ in the points $(0,0)$, (mv_0^+, v_0^+), and $(u_0^+, u_0^+/x^+)$. By Theorem 10, $\mathscr{A}_g$ is convex in the given parameter range and thus it contains the triangle $\mathscr{T}$ with these vertices. It is easy to calculate that the area of this triangular is equal to $u^+\, v^+ (1 - m/x^+)/2$ and we therefore get the bound

$$\text{rejection constant} \le \frac{\text{area of } \mathscr{R}}{\text{area of } \mathscr{T}} = \frac{u_0^+\, v_0^+}{u_0^+\, v_0^+\left(1 - \frac{m}{x^+}\right)/2} = \frac{2}{1 - m/x_0^+} .$$

For an upper bound for the rejection constant it is therefore enough to find an upper bound for m/x_0^+ which is always smaller than one. We have

$$\frac{m(\lambda,\beta)}{x_0^+(\lambda,\beta)} = \frac{\lambda + \sqrt{\beta^2 + (1-\lambda)^2} - 1}{\lambda + \sqrt{\beta^2 + (1+\lambda)^2} + 1} \tag{7}$$

and a straightforward computation shows that the partial derivatives of (7) with respect to β and λ are always positive in the given parameter range. Thus (7) is maximized for $\beta = 1$ and $\lambda = 1$ and we find

$$\frac{m(\lambda,\beta)}{x_0^+(\lambda,\beta)} \le \frac{m(1,1)}{x_0^+(1,1)} = \frac{1}{2+\sqrt{5}}$$

and thus

$$\text{rejection constant} \le \frac{2}{1 - m(1,1)/x_0^+(1,1)} = \frac{2}{1 - \frac{1}{2+\sqrt{5}}} < 2.619$$

as claimed. □

We finally remark, that the proof of Lemma 11 presented in the appendix works completely analogously for this version of the Ratio-of-Uniforms method. We thus know that the limit of the rejection constant for β tending to zero diverges for $\lambda < 1$.

6 A Universally Bounded Hat

In this section we present a hat function for a rejection algorithm when $0 \le \lambda \le 1$ and $\beta \le \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$. For this hat we partition the domain $(0,\infty)$ of g into three intervals, $(0,x_0)$, $(x_0, \max\{x_0, 2/\beta\})$, and $(\max\{x_0, 2/\beta\}, \infty)$, where x_0 is defined in (4). Notice that the second interval might be empty.

Lemma 13 *Let $0 \le \lambda < 1$ and $\beta < 1/2$. Let $x_* = \max\{x_0, \frac{2}{\beta}\}$ and*

$$h_3(x) = x_*^{\lambda-1}\, e^{-\frac{\beta}{2}x} \quad \textit{and} \quad s_3(x) = g(x_*) \cdot e^{r\,(x-x_*)}$$

where

$$r = (\lambda-1)\frac{1}{x_*} - \frac{\beta}{2} + \frac{\beta}{2}\frac{1}{x_*^2}\,.$$

Then $s_3(x) \le g(x) \le h_3(x)$ for all $x \ge x_*$. Moreover,

$$A_3 = \int_{x_*}^{\infty} h_3(x)dx = \frac{2}{\beta}\, x_*^{\lambda-1}\, e^{-\frac{\beta}{2}x_*} \quad \textit{and} \quad \rho_3 = \frac{\int_{x_*}^{\infty} h_3(x)dx}{\int_{x_*}^{\infty} s_3(x)dx} < 2\,e^{\frac{1}{16}} < 2.129\,.$$

Proof Notice that $e^{-\frac{\beta}{2}x} \le 1$ and $x^{\lambda-1} \le x_*^{\lambda-1}$ for all $x \ge x_*$ as $x^{\lambda-1}$ is monotonically decreasing in x. Thus we find

$$g(x) = x^{\lambda-1}\, e^{-\frac{\beta}{2}x}\, e^{-\frac{\beta}{2}\frac{1}{x}} \le x_*^{\lambda-1}\, e^{-\frac{\beta}{2}x} = h_3(x)\,.$$

Now notice that r is the first derivative of $\log(g(x))$ at $x = x_*$ and hence $\tilde{s}_3(x) = \log(g(x_*)) + r\,(x - x_*)$ is the tangent to the log-density at a. Since g is log-convex for $x \ge x_0$ by Lemma 6, $\tilde{s}_3(x) \le \log(g(x))$ and consequently $s_3(x) = \exp(\tilde{s}_3(x)) \le g(x)$ for all $x \ge x_0$ as proposed.

For the second statement notice that $s_3(x)$ is strictly monotonically decreasing, i.e., $r < 0$. Consequently we find

$$A_s = \int_{x_*}^{\infty} s_3(x)dx = g(x_*)\int_{x_*}^{\infty} e^{r(x-x_*)}dx = -\frac{1}{r}\, x_*^{\lambda-1}\, e^{-\frac{\beta}{2}x_*}\, e^{-\frac{\beta}{2}\frac{1}{x_*}}$$

and

$$A_3 = A_h = \int_{x_*}^{\infty} h_3(x)dx = x_*^{\lambda-1}\int_{x_*}^{\infty} e^{-\frac{\beta}{2}x}dx = x_*^{\lambda-1}\,\frac{2}{\beta}\, e^{-\frac{\beta}{2}x_*}$$

and hence

$$\rho_3 = \frac{A_h}{A_s} = \frac{x_*^{\lambda-1}\,\frac{2}{\beta}\, e^{-\frac{\beta}{2}x_*}}{-\frac{1}{r}\, x_*^{\lambda-1}\, e^{-\frac{\beta}{2}x_*}\, e^{-\frac{\beta}{2}\frac{1}{x_*}}} = -\frac{2}{\beta}\, r\, e^{\frac{\beta}{2x_*}} = \left[(1-\lambda)\frac{2}{\beta}\frac{1}{x_*} - \frac{1}{x_*^2} + 1\right] e^{\frac{\beta}{2x_*}}\,.$$

If $x_* = \frac{2}{\beta} \ge x_0$ we find

$$\rho_3 = \left[(1-\lambda)\frac{2}{\beta}\frac{\beta}{2} - \frac{\beta^2}{4} + 1\right] e^{\frac{\beta^2}{4}} = \left[2 - \lambda - \frac{\beta^2}{4}\right] e^{\frac{\beta^2}{4}} < 2\,e^{\frac{1}{16}}$$

where the last inequality follows from our assumption that $\beta < 1/2$.
If $x_* = x_0 = \frac{\beta}{1-\lambda} > \frac{2}{\beta}$, then $1 - \lambda < \frac{\beta^2}{2}$ and we find

$$\rho_3 = \left(\frac{(1-\lambda)^2}{\beta^2} + 1\right) e^{\frac{1-\lambda}{2}} < \left(1 + \frac{\beta^2}{4}\right) e^{\frac{\beta^2}{4}} < \left(1 + \frac{1}{16}\right) e^{\frac{1}{16}}\,.$$

This completes the proof. □

Lemma 14 *Let $0 \le \lambda < 1$ and $\beta < 1/2$ and assume that $x_0 < 2/\beta$. Let*

$$h_2(x) = x^{\lambda-1}\, e^{-\beta} \quad \textit{and} \quad s_2(x) = x^{\lambda-1}\, e^{-1-\frac{\beta^2}{4}} \ .$$

Then $s_2(x) \le g(x) \le h_2(x)$ for all $x \in [x_0, 2/\beta]$. Moreover,

$$A_2 = \int_{x_0}^{\frac{2}{\beta}} h_2(x)dx = \frac{e^{-\beta}}{\lambda}\left((2/\beta)^{\lambda} - x_0^{\lambda}\right) \quad \textit{and} \quad \rho_2 = \frac{\int_{x_0}^{\frac{2}{\beta}} h_2(x)dx}{\int_{x_0}^{\frac{2}{\beta}} s_2(x)dx} < e \ .$$

Proof Let $u(x) = -\frac{\beta}{2}\left(x + \frac{1}{x}\right)$. Then a direct computation shows that $u(x)$ is concave for $x > 0$ and has a global maximum in $x = 1$. As $u(1) = -\beta$ we find

$$g(x) = x^{\lambda-1}\, e^{-\frac{\beta}{2}(x+1/x)} \le x^{\lambda-1}\, e^{-\beta} = h_2(x)\,.$$

On the other hand, $u(x)$ attains its minimum in $[x_0, 2/\beta]$ in one of the boundary points. As

$$u(2/\beta) - u(x_0) = \frac{1+\lambda}{4(1-\lambda)}\left(\beta^2 - 2(1-\lambda)\right) < 0$$

we find $u(2/\beta) < u(x_0)$. The last factor $(\beta^2 - 2(1-\lambda))$ is negative since $x_0 < 2/\beta$. As $u(2/\beta) = -1 - \frac{\beta^2}{4}$ we find for all $x \in [x_0, 2/\beta]$

$$g(x) = x^{\lambda-1}\, e^{-\frac{\beta}{2}(x+1/x)} \ge x^{\lambda-1}\, e^{-1-\frac{\beta^2}{4}} = s_2(x)$$

as claimed. For the second statement we find

$$\rho_2 = \frac{\int_{x_0}^{\frac{2}{\beta}} h_2(x)dx}{\int_{x_0}^{\frac{2}{\beta}} s_2(x)dx} = \frac{\int_{x_0}^{\frac{2}{\beta}} x^{\lambda-1}\, e^{-\beta}\, dx}{\int_{x_0}^{\frac{2}{\beta}} x^{\lambda-1}\, e^{-1-\frac{\beta^2}{4}}\, dx} = \frac{e^{-\beta}}{e^{-1-\frac{\beta^2}{4}}} = e^{1-\beta+\frac{\beta^2}{4}} < e$$

since $0 < \beta < 1/2$. This completes the proof. □

The quasi-density g is log-concave in the remaining interval $(0, x_0)$. Therefore we could use any method that is suitable for constructing a hat function for log-concave densities, e.g., the Ratio-of-Uniforms method as described in Sect. 5.

However, using a constant hat in this interval has an unrivaled cheap setup. So this is our first choice provided that the resulting rejection constant is not too large. It is important to consider that when λ is tending from left to 1, then x_0 is tending to infinity. Thus we cannot find a constant hat with uniformly bounded rejection constant for the parameter range $0 \le \lambda < 1$. Nevertheless, this is possible if we further restrict this range to $0 < \beta < \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$. Indeed we have the following result.

Lemma 15 *Let $0 \le \lambda < 1$ and $0 < \beta \le \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$. Let*

$$h_1(x) = g(m) \quad \textit{and} \quad s_1(x) = g(x_0) \textit{ for } x \ge m \textit{ and } s_1(x) = 0 \textit{ otherwise.}$$

Then $s_1(x) \le g(x) \le h_1(x)$ for all $x \in [0, x_0]$. Moreover,

$$A_1 = \int_0^{x_0} h_1(x)dx = x_0 g(m) \quad \textit{and} \quad \rho_1 = \frac{\int_0^{x_0} h_1(x)dx}{\int_0^{x_0} s_1(x)dx} < 2.72604\ .$$

Proof As g is unimodal with mode m the first inequality immediately follows. For the second statement observe that

$$\rho_1 = \frac{\int_0^{x_0} h_1(x)dx}{\int_0^{x_0} s_1(x)dx} = \frac{g(m)\,x_0}{g(x_0)\,(x_0-m)} = \frac{g(m)}{g(x_0)}\,\frac{1}{1-\frac{m}{x_0}}\;.$$

We therefore need an upper bound for $\frac{g(m)}{g(x_0)}$ and for $\frac{m}{x_0}$. The latter is easily obtained. Using (2) and (4) we find

$$\frac{m}{x_0} = \frac{1-\lambda}{\sqrt{(1-\lambda)^2+\beta^2}+(1-\lambda)} \le \frac{1-\lambda}{(1-\lambda)+(1-\lambda)} = \frac{1}{2}\;. \tag{8}$$

For $\frac{g(m)}{g(x_0)}$ we find

$$\log\left(\frac{g(m(\lambda,\beta))}{g(x_0(\lambda,\beta))}\right) = \frac{\beta^2}{2(1-\lambda)} - \frac{\beta^2}{q(\beta)} - \frac{1}{2}(1-\lambda) - (1-\lambda)\log\left(\frac{1-\lambda}{q(\beta)}\right) \tag{9}$$

where $q(\beta) = \sqrt{\beta^2+(1-\lambda)^2}+(1-\lambda)$. As $q(\beta)$ and hence (9) are monotonically increasing for $\beta \ge 0$ we find that (9) attains its maximum on the curve defined by $\beta = \frac{2}{3}\sqrt{1-\lambda}$. Just plugging in that value of β into (9) we obtain

$$p(\lambda) := \log\left(\frac{g(m(\lambda,\frac{2}{3}\sqrt{1-\lambda}))}{g(x_0(\lambda,\frac{2}{3}\sqrt{1-\lambda}))}\right)\;.$$

A tedious straightforward computation then gives for the second derivative

$$p''(\lambda) = \frac{4}{3\sqrt{13-22\lambda+9\lambda^2}\left(3-3\lambda+\sqrt{13-22\lambda+9\lambda^2}\right)^2}$$

which is non-negative for all $\lambda \in [0,1]$. Hence $p(\lambda)$ attains its maximum on the boundary of that domain and we find $p(\lambda) \le p(0) < 0.3097$.

Consequently, $\frac{g(m)}{g(x_0)} < \exp(0.3097) < 1.36302$ and thus

$$\rho_3 = \frac{g(m)}{g(x_0)}\,\frac{1}{1-\frac{m}{x_0}} < \frac{1.36302}{1-\frac{1}{2}} = 2.72604$$

as claimed. □

When we combine the results from Lemmata 13, 14, and 15 we obtain the following hat for a rejection method. Figure 3 shows an example of this hat function and the squeezes.

Theorem 16 *Let* $x_* = \max\{x_0, \frac{2}{\beta}\}$ *and*

$$h(x) = \begin{cases} 0 & \text{for } x \le 0,\\ g(m) & \text{for } 0 < x \le x_0,\\ x^{\lambda-1}\,e^{-\beta} & \text{for } x_0 < x < 2/\beta,\\ x_*^{\lambda-1}\,e^{-\frac{\beta}{2}x} & \text{for } x \ge \max\{x_0, 2/\beta\}. \end{cases}$$

If $0 \le \lambda < 1$ *and* $0 < \beta \le \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$ *then* $h(x) \ge g(x)$ *for all* $x \ge 0$. *Moreover,*

$$\frac{\int_0^\infty h(x)dx}{\int_0^\infty g(x)dx} \le \max\{\rho_1,\rho_2,\rho_3\} < 2.72604$$

that is, the rejection constant is uniformly bounded from above by 2.72604.

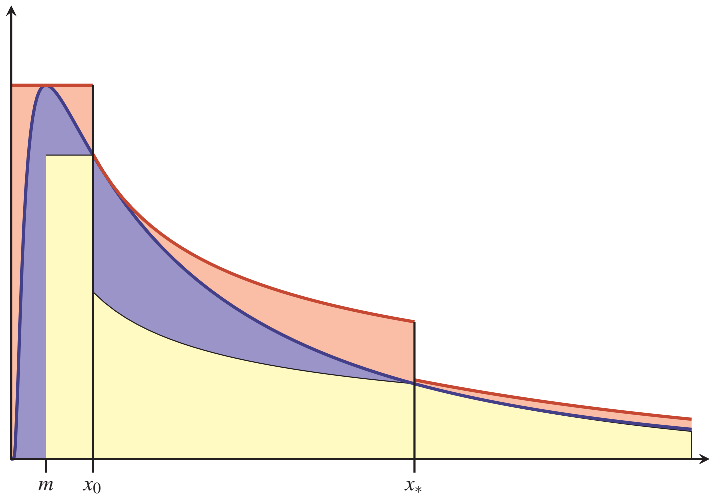

Fig. 3 Hat function of Thm. 16. The squeeze is used in Lemmata 13, 14, and 15 to estimate an upper bound for the rejection constant. ($\lambda = 0.5, \beta = 0.45$)

7 The Algorithm

In this section we compile a generator for GIG distributed random variates when $0 \le \lambda \le 1$ and $0 < \beta \le 1$ using our results from Sections 5 and 6. Algorithm 1 compiles the details for the rejection method based on the hat from Section 6.

Remark 17 Notice that the formula for computing A_2 in Lemma 14 does not work when $\lambda = 0$. However, by l'Hospital's rule we then find

$$A_2 = \lim_{\lambda \to 0} k_2 \left((2/\beta)^\lambda - x_0^\lambda \right) / \lambda = k_2 \log \left(\frac{2}{\beta x_0} \right) .$$

Similarly, we have to use the corresponding limit in Step 16 of Algorithm 1.

For the sake of self-containedness we also state the Ratio-of-Uniforms method without mode shift in Algorithm 2 and the Ratio-of-Uniforms method with mode shift in Algorithm 3. For the latter we used Cardano's formula to find the minimal bounding rectangle.

The combination of these two (three) algorithms results in a generator for the GIG distribution with uniformly bounded performance. We remark here that numerical experiments have shown that the observed rejection constants are always smaller than our estimated upper bounds. Indeed we found that it never exceeded 1.5. Figure 4 shows the rejection constant for parameter range $0 \le \lambda \le 1.5$ and $0 < \beta \le 1.5$. The parameter regions of the three algorithms can be easily seen.

8 Conclusion

We have characterized the performance of the very popular random variate generation algorithm for the GIG distribution suggested by Dagpunar (1989). Considering the concavity property of the transformed density we have shown that for $\lambda \ge 1$ or $\beta \ge \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$

Algorithm 1 Rejection method for non-$T_{-1/2}$-concave part

Input: Parameters λ, β with $0 \le \lambda < 1$ and $0 < \beta \le \frac{2}{3}\sqrt{1-\lambda}$.
Output: GIG distributed random variate X.

▷ Setup

$m \leftarrow \beta / \left((1-\lambda) + \sqrt{(1-\lambda)^2 + \beta^2}\right)$
$x_0 \leftarrow \beta/(1-\lambda), \quad x_* \leftarrow \max\{x_0, 2/\beta\}$
$k_1 \leftarrow g(m), \quad A_1 \leftarrow k_1 x_0$ ▷ subdomain $(0, x_0)$
if $x_0 < 2/\beta$ **then**
    $k_2 \leftarrow e^{-\beta}, \quad A_2 \leftarrow k_2 \left((2/\beta)^\lambda - x_0^\lambda\right)/\lambda$ ▷ subdomain $(x_0, 2/\beta)$
    [if $\lambda = 0$ then $A_2 \leftarrow k_2 \log(2/\beta^2)$]
else
    $k_2 \leftarrow 0, \quad A_2 \leftarrow 0$
$k_3 \leftarrow x_*^{\lambda-1}, \quad A_3 \leftarrow 2k_3 \exp(-x_*\beta/2)/\beta$ ▷ subdomain (x_*, ∞)
$A \leftarrow A_1 + A_2 + A_3$

▷ Generator

repeat
    generate $U \sim \mathscr{U}(0,1)$ and $V \sim \mathscr{U}(0,A)$
    if $V \le A_1$ **then** ▷ subdomain $(0, x_0)$
        $X \leftarrow x_0 V/A_1, \quad h \leftarrow k_1$
    else if $V \le A_1 + A_2$ **then** ▷ subdomain $(x_0, 2/\beta)$
        $V \leftarrow V - A_1$
        $X \leftarrow \left(x_0^\lambda + V\lambda/k_2\right)^{1/\lambda}, \quad h \leftarrow k_2 X^{\lambda-1}$
        [if $\lambda = 0$ then $X \leftarrow \beta \exp(V \exp(\beta))$]
    else ▷ subdomain (x_*, ∞)
        $V \leftarrow V - (A_1 + A_2)$
        $X \leftarrow -2/\beta \log\left(\exp(-x_*\beta/2) - V\beta/(2k_3)\right), \quad h \leftarrow k_3 \exp(-X\beta/2)$
until $Uh \le g(X)$
return X

Algorithm 2 Ratio-of-Uniforms without mode shift

Input: Parameters λ, β with $0 \le \lambda \le 1$ and $\min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\} \le \beta \le 1$.
Output: GIG distributed random variate X.

▷ Setup: Compute minimal bounding rectangle

$m \leftarrow \beta / \left((1-\lambda) + \sqrt{(1-\lambda)^2 + \beta^2}\right)$
$x^+ \leftarrow \left((1+\lambda) + \sqrt{(1+\lambda)^2 + \beta^2}\right)/\beta$
$v^+ \leftarrow \sqrt{g(m)}$
$u^+ \leftarrow x^+ \sqrt{g(x^+)}$

▷ Generator

repeat
    Generate $U \sim \mathscr{U}(0, u^+)$ and $V \sim \mathscr{U}(0, v^+)$
    $X \leftarrow U/V$
until $V^2 \le g(X)$
return X

the rejection constant of this Ratio-of-Uniforms algorithm with mean shift is bounded by two. In addition we have shown that for $\lambda < 1$ and β tending to zero the rejection constant diverges. We also have observed that for $\lambda < 1$ and $\beta < 10^{-8}$ the implementation of that algorithm available on the internet returns random variates of a wrong distribution due to rounding errors when calculating the minimal bounding rectangle.

Finally a new, simple rejection algorithm has been presented. We have also proven that it has a uniformly bounded rejection constant for $\lambda < 1$ and $\beta < \min\{\frac{1}{2}, \frac{2}{3}\sqrt{1-\lambda}\}$. The

Algorithm 3 Ratio-of-Uniforms with mode shift (Dagpunar-Lehner)

Input: Parameters λ, β with $\lambda > 1$ and $\beta > 1$.
Output: GIG distributed random variate X.

▷ Setup: Compute minimal bounding rectangle

$m \leftarrow \left(\sqrt{(\lambda-1)^2+\beta^2}+(\lambda-1)\right)/\beta$

$a \leftarrow -\frac{2(\lambda+1)}{\beta} - m, \quad b \leftarrow \frac{2(\lambda-1)}{\beta} m - 1, \quad c \leftarrow m$ ▷ Find solution of (5)

$p \leftarrow b - \frac{a^2}{3}, \quad q \leftarrow \frac{2a^3}{27} - \frac{ab}{3} + c$ ▷ Coefficient of reduced form

$\phi \leftarrow \arccos\left(-\frac{q}{2}\sqrt{-\frac{27}{p^3}}\right)$ ▷ Cardano's formula

$x^- \leftarrow \sqrt{-\frac{4}{3}p} \cdot \cos\left(\frac{\phi}{3}+\frac{4}{3}\pi\right) - \frac{a}{3}$

$x^+ \leftarrow \sqrt{-\frac{4}{3}p} \cdot \cos\left(\frac{\phi}{3}\right) - \frac{a}{3}$

$v^+ \leftarrow \sqrt{g(m)}$

$u^- \leftarrow (x^- - m)\sqrt{g(x^-)}, \quad u^+ \leftarrow (x^+ - m)\sqrt{g(x^+)}$

▷ Generator

repeat
    Generate $U \sim \mathcal{U}(u^-, u^+)$ and $V \sim \mathcal{U}(0, v^+)$
    $X \leftarrow U/V + m$
until $V^2 \le g(X)$
return X

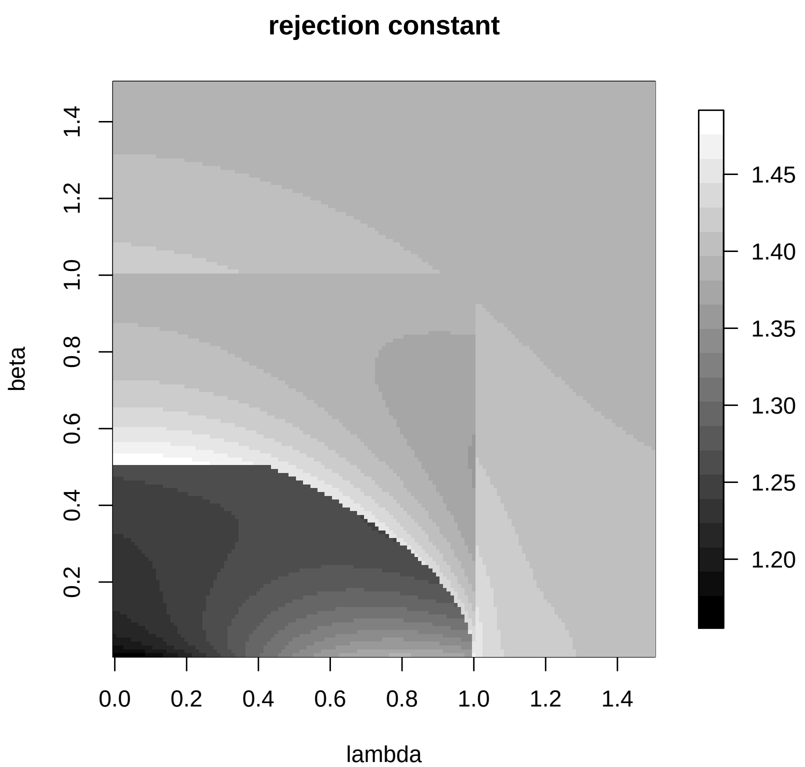

Fig. 4 Rejection constants for the proposed combined algorithm.

combination of the two algorithms thus results in a generator for the GIG distribution with uniformly bounded performance.

Acknowledgments

The authors gratefully acknowledge the useful suggestions of the area editor and two anonymous referees that helped to improve the presentation of the paper.

References

Abramowitz M, Stegun IA (eds) (1972) Handbook of mathematical functions, 9th edn. Dover, New York

Ahrens JH (1995) A one-table method for sampling from continuous and discrete distributions. Computing 54(2):127–146

Atkinson AC (1982) The simulation of generalized inverse Gaussian and hyperbolic random variables. SIAM J Sci Statist Comput 3:502–515

Barndorff-Nielsen O, Blæsild P, Halgreen C (1978) First hitting time models for the generalized inverse gaussian distribution. Stochastic Processes and their Applications 7(1):49–54, DOI 10.1016/0304-4149(78)90036-4

Dagpunar J (1988) Principles of Random Variate Generation. Clarendon Oxford Science Publications, Oxford, U.K.

Dagpunar JS (1989) An easily implemented generalised inverse Gaussian generator. Comm Statist B — Simulation Comput 18:703—710

Dagpunar JS (2007) Simulation and Monte Carlo with Applications in Finance and MCMC. Wiley

Derflinger G, Hörmann W, Leydold J (2010) Random variate generation by numerical inversion when only the density is known. ACM Trans Model Comput Simul To appear. Preprint available as Research Report Series of the Department of Statistics and Mathematics at WU Vienna, No. 90, June 2009. http://epub.wu.ac.at/, oai:epub.wu-wien.ac.at:epub-wu-01_f41.

Devroye L (1984) A simple algorithm for generating random variates with a log-concave density. Computing 33(3–4):247–257

Devroye L (1986) Non-Uniform Random Variate Generation. Springer-Verlag, New-York

Eberlein E, Keller U (1995) Hyperbolic distributions in finance. Bernoulli 1:281–299

Hörmann W (1995) A rejection technique for sampling from T-concave distributions. ACM Trans Math Software 21(2):182–193

Hörmann W, Leydold J, Derflinger G (2004) Automatic Nonuniform Random Variate Generation. Springer-Verlag, Berlin Heidelberg

Johnson N, Kotz S, Balakrishnan N (1994) Continuous Univariate Distributions, Volume 1, 2nd edn. Wiley, New York

Jørgensen B (1982) Statistical Properties of the Generalized Inverse Gaussian Distribution, Lecture Notes in Statistics, vol 9. Springer-Verlag, New York-Berlin

Kinderman AJ, Monahan JF (1977) Computer generation of random variables using the ratio of uniform deviates. ACM Trans Math Software 3(3):257–260

Lehner K (1989) Erzeugung von Zufallszahlen aus zwei exotischen Verteilungen. Diploma thesis, Tech. Univ. Graz, Austria, 107 pp.

Leydold J (2000) Automatic sampling with the ratio-of-uniforms method. ACM Trans Math Software 26(1):78–98, DOI http://doi.acm.org/10.1145/347837.347863

Leydold J (2001) A simple universal generator for continuous and discrete univariate T-concave distributions. ACM Trans Math Software 27(1):66–82

Leydold J, Hörmann W (2011) Generating generalized inverse gaussian random variates by fast inversion. Computational Statistics and Data Analysis 55(1):213–217, DOI 10.1016/j.csda.2010.07.011

Seshadri V (1999) The inverse Gaussian distribution, Lecture Notes in Statistics, vol 137. Springer, New York

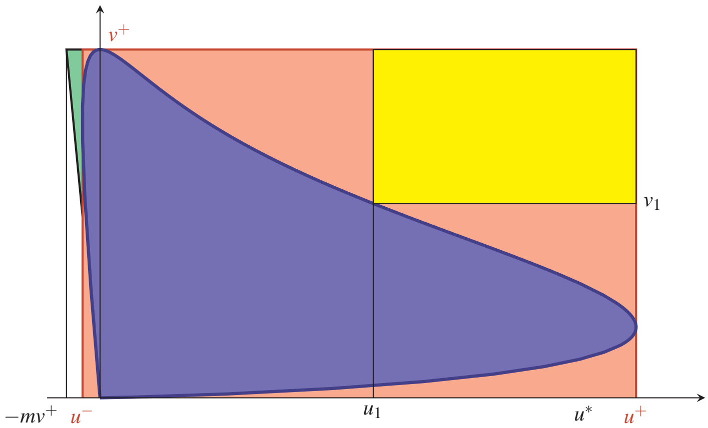

Fig. 5 $\mathscr{A}_g$, $\mathscr{R}$, $(-mv^+, u+) \times (0, v^+)$, and (u_1, v_1) in proof of Lemma 11.

Appendix

Proof of Lemma 11

Let $\lambda \in [0,1)$ be fixed. Recall that $\mathscr{R} = (u^-, u^+) \times (0, v^+)$. Notice that $\mathscr{A}_g$ does not contain any point left of the line $u = -mv$ since we have used $\mu = m$ in (3), i.e., we have shifted quasi-density g by the mode m to the left, see Fig. 5. Consequently, $-mv^+ \leq u^-$ and $\mathscr{A}_g \subseteq (-mv^+, u+) \times (0, v^+)$. Let x^+ defined as in Sect. 2, i.e., it is the unique root of (5) greater than m. Thus $(x-m)\sqrt{g(x)}$ is monotonically increasing in $[m, x^+]$.

Now choose $x_1 \in (0, x^+ - m)$ and let (u_1, v_1) be the point on the boundary of $\mathscr{A}_g$ corresponding to x_1, i.e., $v_1 = \sqrt{g(x_1+m)}$ and $u_1 = xv_1$. Then $\mathscr{A}_g$ does not intersect the open rectangle $(u_1, u^+) \times (v_1, v^+)$ and thus

$$\mathscr{A}_g \subseteq (-mv^+, u^+) \times (0, v^+) \setminus (u_1, u^+) \times (v_1, v^+) .$$

We therefore find

$$\frac{|\mathscr{A}_g|}{|\mathscr{R}|} \leq \frac{(mv^+)v^+ + u^+ v_1 + u_1 v^+ - u_1 v_1}{(u^+ - u^-)v^+} \leq m\frac{v^+}{u^+} + \frac{v_1}{v^+} + \frac{u_1}{u^+} . \tag{10}$$

Now let

$$x_0^+(\beta) = \frac{1}{\beta}\left((1+\lambda) + \sqrt{(1+\lambda)^2 + \beta^2}\right) .$$

It is the unique maximum of $x\sqrt{g(x)}$, see Sect. 5. Since $\left((x-m)\sqrt{g(x)}\right)' \geq \left(x\sqrt{g(x)}\right)'$ for all $x \geq m$, we find $x_0^+ \leq x^+$. Now define

$$u^* = (x_0^+ - m)\sqrt{g(x_0^+)} .$$

Clearly $u^* \leq u^+ = \sup(x-m)\sqrt{g(x)}$, and thus $\varepsilon = u^+ - u^* \geq 0$. From (10) we then obtain

$$\frac{|\mathscr{A}_g|}{|\mathscr{R}|} \leq m\frac{v^+}{u^*} + \frac{v_1}{v^+} + \frac{u_1}{u^*} . \tag{11}$$

Now set

$$x_1(\beta) = (x_0^+(\beta))^\beta, \quad v_1(\beta) = \sqrt{g(x_1(\beta)+m(\beta))}, \quad \text{and} \quad u_1(\beta) = x_1(\beta)v_1(\beta)\,.$$

We first have to check whether the condition $x_1 \in (0, x^+ - m)$ is fulfilled. For the limits $\beta \to 0$ we find

$$\begin{aligned}
\lim_{\beta\to 0} x_0^+(\beta) &= \lim_{\beta\to 0} \frac{1}{\beta}\left((1+\lambda)+\sqrt{(1+\lambda)^2+\beta^2}\right) = +\infty \\
\lim_{\beta\to 0} x_1(\beta) &= \lim_{\beta\to 0}(x_0^+(\beta))^\beta = \lim_{\beta\to 0} 2(\lambda+1)^\beta\left(\frac{1}{\beta}\right)^\beta = 1 \\
\lim_{\beta\to 0} m(\beta) &= \lim_{\beta\to 0} \frac{\beta}{\sqrt{(1-\lambda)^2+\beta^2}+(1-\lambda)} = 0\,.
\end{aligned}$$

An immediate consequence is that for sufficiently small $\beta > 0$, $x_1(\beta) < x_0^+(\beta) - m(\beta) \le x^+(\beta) - m(\beta)$ which shows that $x_1(\beta) \in (0, x^+ - m)$ when β is close enough to zero. Thus inequality (11) holds. Moreover,

$$\begin{aligned}
\lim_{\beta\to 0} v_1(\beta)^2 &= \lim_{\beta\to 0} g(x_1(\beta)+m(\lambda,\beta)) \\
&= \lim_{\beta\to 0}(x_1+m)^{\lambda-1}\cdot \lim_{\beta\to 0}\exp\left(-\frac{\beta}{2}\left(\frac{1}{x_1+m}+(x_1+m)\right)\right) = 1 \\
\lim_{\beta\to 0} u_1(\beta) &= \lim_{\beta\to 0} v_1(\beta)\cdot \lim_{\beta\to 0} x_1(\beta) = 1\,.
\end{aligned}$$

For the denominators in (11) we find

$$\begin{aligned}
\lim_{\beta\to 0}(v^+(\beta))^2 &= \lim_{\beta\to 0} m^{\lambda-1}\cdot \lim_{\beta\to 0}\exp\left(-\frac{\beta}{2}\left(\frac{1}{m}+m\right)\right) = \lim_{\beta\to 0} m^{\lambda-1}\cdot\exp(\lambda-1) = +\infty \\
\lim_{\beta\to 0}(u^*(\beta))^2 &= \lim_{\beta\to 0}(x_0^+(\beta)-m(\beta))^2\, g(x_0^+(\beta)) \\
&= \lim_{\beta\to 0}(x_0^+-m)^2(x_0^+)^{\lambda-1}\cdot \lim_{\beta\to 0}\exp\left(-\frac{\beta}{2}\left(\frac{1}{x_0^+}+x_0^+\right)\right) \\
&= \lim_{\beta\to 0}(x_0^+)^{\lambda+1}\cdot\exp(1+\lambda) = +\infty\,.
\end{aligned}$$

Finally,

$$\lim_{\beta\to 0} m(\beta)\frac{v^+(\beta)}{u^*(\beta)} = \lim_{\beta\to 0}\frac{m^\lambda}{(x_0^+)^{\lambda+1}}e^{-2} = \lim_{\beta\to 0} const\cdot\frac{(\beta)^\lambda}{(1/\beta)^{\lambda+1}} = \lim_{\beta\to 0}(\beta)^{2\lambda+1} = 0\,.$$

Collecting all limits we find that all fractions on the right hand side of inequality (11) converge to 0 and thus $\lim_{\beta\to 0}\frac{|\mathscr{A}_g(\beta)|}{|\mathscr{R}(\beta)|} = 0$ as claimed. $\square$